

CNN を用いた手書き文字認識と印字システムの設計

永田研究室 F117039 DANG QUANG TIEN

1. 目的

手書き文字の認識は、運輸・郵便業界で実用化されているが、手書き文字の中には癖や文字の濃さによって人間でも判別に困るものが含まれていることもあり、完璧な識別を行うことは非常に困難である。また、認識する文字の種類が増加するにつれて類似した形状を持つものも増加し、それによっても誤認識が増加する。こうした類似の画像を分類する課題の一つとして一般画像認識がある。これは、風景を撮影した画像に写る物体を決められたカテゴリに正確に分類することを目標とする。最近、この分野で認識精度を飛躍的に高めたとして注目された技術が畳み込みニューラルネットワーク(CNN)である。そこで、本研究ではCNNを用いて手書き文字の認識を試みる。その際、文字の種類が増加すると、似た形状の文字を含む画像の誤分類が増加することが考えられるため、アルファベットの大文字と小文字をまとめて大文字として認識する。また、文字認識を応用して、ロボットに認識した文字を描画させることも試みる。

2. 研究内容

提案するシステムは、WebカメラとDobot社製のロボットアームであるDobot Magician、および制御用のPCで構成され、印字対象であるホワイトボードを含め、図1のように配置されている。文字認識は、畳み込みニューラルネットワーク(CNN)を用いてホワイトボード上に書かれた文字を、大文字のA~Zの26カテゴリに分類することで行う。今回使用するCNNは、Matlab上で公開されているImageNet Large Scale Visual Recognition Challengeで事前に学習したAlexNetと呼ばれるCNNをベースに、次の2つのプロセスを経て作成する。まず、AlexNetの出力層を含む3層を26カテゴリ分類用にして新たなCNNを再設計する。次に、米国国立標準技術研究所で公開されている文字認識用画像データセットでこのCNNを再学習する。さらに、手書き文字が書かれている領域をカメラ画像から自動的に抽出するために、次のような画像処理を行う。まず、撮影した画像を二値化し、その中で最も大きな面積を持つ領域を文字と推定する。次に、文字を構成するピクセルの位置情報から文字の図心を計算する。最後に文字が写る領域を図心から正方形に切り取る。また、ロボットを用いて文字を書くために、文字の形状に合ったペン先の軌道を生成する。そこで、Adobe Illustratorを用いて大文字のA~Z、書体はCalibriのDFX形式で文字の形状を保存し、CAD/CAMソフトのCreoで文字の加工のシミュレーションを行い、目標経路となるCLSデータを作成する。

3. 結果

研究室の5人のメンバーの手書き文字を使って印字実験を行った。その内、ある程度薄い線の場合には、二値化処理に使用する閾値を手動で変更することにより、ホワイトボードと文字とを区別することができ、正確に文字を認識することができた。しかし、文字の大きさや癖、形状の類似度など(例えば、大文字のIと小文字のl, qとgなど)の要因により、別の文字として誤認識される場合もあった。今後はデータセットを増やすことにより認識精度を向上させたいと考えている。

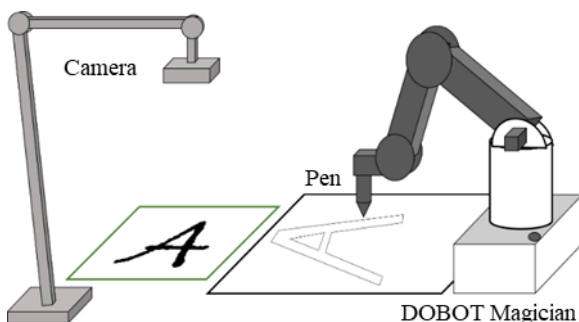


Fig. 1 System configuration.

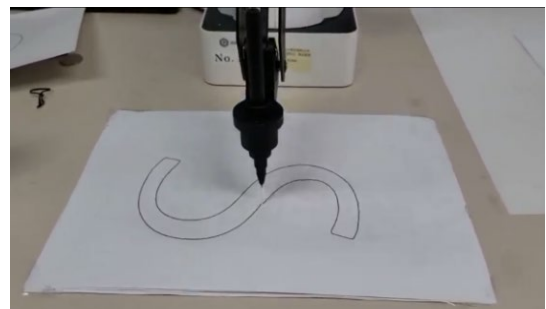


Fig. 2 Experimental result.